

## 冷冻电镜技术在表观遗传学研究中的应用\*

李承珉<sup>1, 2)\*\*</sup> 周旻<sup>1, 2)\*\*</sup> 李晓敏<sup>1, 2)</sup> 朱平<sup>1)\*\*\*</sup>

(<sup>1)</sup> 中国科学院生物物理研究所, 生物大分子国家重点实验室, 北京 100101; (<sup>2)</sup> 中国科学院大学, 北京 100049)

**摘要** 染色质装配、修饰和重塑复合体, 以及它们和核小体、染色质等一起形成的超大分子复合体的精细结构解析, 对于在原子水平揭示表观遗传信息建立、维持和调控的分子机制至关重要。近年来, 迅速发展的冷冻电镜三维重构技术对于解析这些多亚基、大分子质量、柔性超大分子复合体的结构带来了很好的机遇。本文综述了冷冻电镜三维重构技术在表观遗传学相关的结构研究领域的一些应用和进展。

**关键词** 表观遗传调控, 冷冻电镜, 核小体, 染色质

**学科分类号** Q6-33, Q34

**DOI:** 10.16476/j.pibb.2015.0295

表观遗传学研究所关注的是, 在 DNA 序列不改变的情况下基因表达发生的可遗传的改变。作为遗传学新的分支, 表观遗传学是近年来生命科学的热门领域之一。在表观遗传调控中, 由 DNA 缠绕在组蛋白八聚体上形成的核小体是一个极为重要的基本环节。一方面, 作为最基本的结构单元, 核小体压缩折叠形成高级结构染色质及染色体的机制方式是我们理解表观遗传的关键; 另一方面, 诸多的表观遗传调控因子都是通过与核小体及其组成的染色质相互作用来行使其丰富的功能, 这也正是表观遗传学能如此精彩的一个原因。

核小体、染色质的结构以及它们如何受各种因子的调控一直是表观遗传学研究者极为关注的内容, 但目前应用最为广泛的结构生物学研究方法——X 射线晶体学在获得这些典型的超大分子复合体结构方面遇到了很大的瓶颈和挑战。电镜技术, 尤其是冷冻电镜技术, 是结构生物学里非常重要也是最近这几年引起研究者极大关注的研究手段。一方面, 它使用电子聚焦, 因此拥有远高于光学显微镜的分辨能力。另一方面, 它不需要样品结晶, 对样品纯度也相对要求较低, 因此研究的门槛远低于 X 射线晶体学技术。虽然目前处理生物学样品时还存在信噪比低等一些劣势, 但是随着直接电子探测器的应用以及更好的图像处理 and 重构算法

的开发, 冷冻电镜的前沿已经进入原子分辨率的时代<sup>[1]</sup>。

在这里, 我们回顾整理了冷冻电镜三维重构技术在表观遗传学相关结构研究领域的一些应用。主要的内容聚焦在两个方面: 核小体相关因子复合物三维结构的研究; 染色质的高级结构尤其是 30 nm 纤维结构的研究。

### 1 核小体相关因子复合物的结构研究

核小体修饰及染色质重塑在表观遗传调控中起着重要作用。核小体相关结构因子的冷冻电镜研究主要关注 3 个方面: 染色质重塑复合物(chromatin remodeling complexes)、组蛋白修饰复合物(histone modification complexes)和染色质结构蛋白(chromatin architectural proteins)<sup>[2]</sup>。大多数情况下, 这些蛋白复合物亚基数众多、分子质量巨大, 而且

\* 国家重点基础研究发展计划(973)(2015CB856200), 中国科学院战略性先导专项(B 类)(XDB08010100), 中国科学院重点部署项目(KJZD-EW-L05)和国家自然科学基金(31230018, 31425007, 31521002)资助项目。

\*\* 共同第一作者

\*\*\* 通讯联系人。

Tel: 010-64888799, E-mail: zhup@ibp.ac.cn

收稿日期: 2015-09-25, 接受日期: 2015-09-28

它们结合核小体的方式高度动态, 很难用结晶的方式来得到高分辨率的结构. 即使对于电镜技术而言, 这些多亚基, 高柔性的复合物结构解析也是一个巨大的挑战. 目前已发表的核小体相关复合物的冷冻电镜结构研究中, 除了极少数研究的三维重构分辨率达到了亚纳米级精度范围外, 大部分复合物的结构解析精度都还在 2 nm 左右. 虽然较低分辨率的电镜密度图不能提供很细节的信息, 但是在 X 射线、核磁共振、质谱、生化分析等其他研究手段的辅助下, 这些结果还是可以帮助人们对核小体相关因子本身的整体结构以及它们与核小体的结合方式有更进一步的了解.

### 1.1 染色质重塑复合物

基因组的转录调控, 是真核生命体内最重要的过程之一. 转录过程需要与转录相关的蛋白(如 RNA 聚合酶、激活子、抑制子)直接作用于编码区域的 DNA 上. 但是, 需要转录的 DNA 很多时候紧密缠绕在组蛋白上, 并以核小体的形式折叠压缩成染色质, 无法直接与转录相关蛋白接触. 因此, 转录过程的完成需要其他蛋白质的帮助, 而 ATP 依赖的染色质重塑复合物就属于其中最主要的一类. 它们在细胞中极其丰富, 几乎每 10 个核小体就有一个 ATP 依赖的重塑复合物. 这一类复合物利用 ATP 水解提供自由能打开组蛋白与 DNA 的连接点, 再重新塑造核小体, 以此来动态调控缠绕在组蛋白上 DNA 的状态, 进而决定基因的表达与细胞的命运. 染色质重塑的方式多种多样: 滑动核小体位置、剔除组蛋白八聚体或二聚体或改变组蛋白成分(即某个组蛋白交换成其变体或者反过来)<sup>[3]</sup>. ATP 依赖的染色质重塑复合物都有一个保守的 ATP 酶(ATPase)核心, 属于 SF2(superfamily 2)家族<sup>[4]</sup>. 依据 ATPase 的结构特征又可以划分为 4 个主要的亚家族: SWI/SNF、ISWI、CHD、INO80. 每个亚家族都有自己独特的结构域分布, 不同的催化活性, 不一样的作用位点与底物. 目前, 除了 CHD 家族外, 其他 3 个家族都有了相应的电镜结构研究.

#### 1.1.1 SWI/SNF 复合物

SWI/SNF(switching defective/sucrose non-fermenting)重塑复合物首先在酵母中发现<sup>[5]</sup>, 一般由 8~14 种不同的亚基组成, 分子质量大多大于 1 Mu, 主要结构特征是含有一个能招募肌动蛋白和肌动相关蛋白的 N 端 HSA 结构域(helicase SANT domain), 以及一个能结合在组蛋白的脱乙酰赖氨酸上的 C 端溴域(bromodomain)<sup>[6]</sup>. 大部分真核生物体内含有 2

种相关的 SWI/SNF 重塑复合物: 在酵母中有 SWI/SNF 和 RSC (remodels the structure of chromatin); 在果蝇中有 BAP(brahma associated protein)和 PBAP (polybromo-associated BAP); 在人类中有 BAF (Brg1 associated factors) 和 PBAF (polybromo-associated BAF)<sup>[7]</sup>. 目前已发表的电镜观察复合物的结构中, 有 2 个 SWI/SNF 相关复合物, 5 个 RSC 相关复合物<sup>[8]</sup>. 这些电镜重构结果展示了一个相似的特征: 中央凹陷的 C 状球形蛋白复合物(图 1a~f).

已发表的 4 个酵母 RSC 复合物以及人类中与 RSC 同源的 PBAF 复合物结构非常相似: 4 个球状的结构域与 1 个在空间上能容纳单个核小体的深凹排列成 C 状. 它们能完全包裹住核小体, 与相应的生化实验相符<sup>[9-13]</sup>. 值得注意的是, 在图 1d 展示的 RSC-核小体的密度图中并不能看到完整的核小体密度, 而只是一部分组蛋白<sup>[10]</sup>, DNA 和某个 H2A/H2B 缺失. 在 RSC 与 PBAF 的重构结果中, 可以看到明显的结构柔性<sup>[12-13]</sup>. 在图 1a~e 中箭头所指的耳垂状结构中, 可以看到两种状态: 一种开放的状态和一种闭合的状态. 当一种能激活 RSC 重塑活性的乙酰化组蛋白 H3 拟肽存在的时候, RSC 会更倾向于闭合的状态.

酵母的 SWI/SNF 复合物的结构看起来和 RSC 有较大差异. 电镜三维重构结果显示, SWI/SNF 的结构上有一个较浅的凹陷(图 1f), 可以正好容纳但是无法包住一个核小体. 这个结论被 DNA 足印法(DNA footprinting)和光敏性特异位点交联(photoreactive site-specific cross-linking)实验所验证, 确实只有一边的核小体 DNA 与 SWI/SNF 发生作用<sup>[14-15]</sup>.

#### 1.1.2 ISWI 复合物

ISWI(imitation switch)家族与 SWI/SNF 相比小了很多, 通常只含有 2~4 个亚基<sup>[16]</sup>. 它们的 ATPase 亚基含有一个 N 端 Snf2 驱动结构域和一个由 HAND、SANT、SLIDE(HSS)组成的负责识别核小体的三元结构域. SANT 结构域在结构上与 c-Myb DNA 结合区域相关, 带负电荷, 因此可能与组蛋白有相互作用, 如未修饰的 H4 尾巴; SLIDE(SANT-like ISWI domain)结构域主要与核小体接近二次轴(dyad axis)的 DNA 作用; HAND 结构域则可能同时与组蛋白和 DNA 结合或识别<sup>[17]</sup>. 在酵母中, ISWI 家族成员包括 ISW1a、ISW1b 和 ISW2. 在人和果蝇中则包括 ACF、CHRAC 和

NURF.

在电镜结构解析以前, 研究者凭借光敏反应交联和足印法已经知道, HSS 结构域和连接 DNA 有相互作用, ATPase 驱动则结合在 SHL + 2 的位置(一个非常关键的位点, 核小体上热力学最不稳定 2 个位点之一)<sup>[18]</sup>.

2008 年报道了这个家族的第一个电镜结构, 人源 CHRAC 复合物的负染结构. 从整体上呈多叶状, 由于看不到明显的核小体结合位置, 关于它和核小体的相互作用只能猜测<sup>[19]</sup>.

2009 年 Narlikar 实验室解析了人源 ACF 复合物中负责催化的亚基 SNF2h 与核小体的负染结构<sup>[20]</sup>. 他们发现 SNF2h 与核小体的结合比例为 2:1, 2 个 SNF2h 分别结合在单个核小体的两端. 虽然生化数据显示 ATPase 结构域结合在 SHL+/-2 的位置以及与 H4 N 端尾巴有相互作用, 但是从密度图上并不能看到连接 DNA 和 HSS 结构域(图 1g). 依据这个密度, 作者提出了一个关于 ISWI 作用机制的模型: 2 个 ISWI 的合作协同驱动, 交替性作用(alternating action), 而不是拔河(tug of war). 在这个模型里, 每一个 ATPase 有序地作用在核小体侧面 DNA 上, 一次只有一个 ATPase 处于活动状态. 在连接 DNA 更长的那端的 ATPase 会更快地水解 ATP, 决定着转位的方向, 此时另一端的 ATPase 提供一个稳定支撑的作用, 直到核小体两端的 DNA 长度变得一样<sup>[20]</sup>.

但是 2011 年 Richmond 实验室提出了另外一种作用机制模型. 他们首先解析了 ISW1a( $\Delta$ ATPase)-Ioc3 的复合物以及其与 DNA 结合的晶体结构, 然后利用冷冻电镜三维重构技术发现, ISW1a( $\Delta$ ATPase)与核小体的作用方式随延伸连接 DNA 的长度而变化. 当核小体只有单个 45 bp 的连接 DNA 时, 2 个拷贝的 ISW1a( $\Delta$ ATPase)可以桥接 2 个核小体. 所以在他们的模型里, ISW1a 作用如同一把空间上的尺, 利用自身的直径形状来“测量”2 个相连核小体的距离(图 1h)<sup>[21]</sup>.

目前, 人们还需要更多的结果及更高的分辨率来分析这两种模型哪个更为准确地反映了 ISWI 复合物的实际作用机制.

### 1.1.3 INO80 复合物

INO80 家族的复合物包括 INO80 (inositol requiring80)和 SWR1 (sick with Rat8), 分子质量都在 Mu 级别, 亚基数量众多, 在 Snf2-ATPase 结构域被插入了一段 RecA2 方式折叠的蛋白, 负责招

募 AAA+家族的 Rvb1/2 螺旋酶<sup>[22]</sup>. INO80 和 SWR1 还共有一个与 SWI/SNF 家族相似的 HSA 结构域, 负责招募 actin 和 Arp 蛋白(actin-related proteins). INO80 和 SWR1 都含有 actin 与 Arp4. 从它们共有的亚基以及相似的一级序列来看, INO80 与 SWR1 应该有高度相似的拓扑结构, 但是目前的电镜研究结果却显示并非如此<sup>[23]</sup>.

2013 年 Karl-Peter Hopfner 研究组首次报道了酿酒酵母(*Saccharomyces cerevisiae*)中 INO80 的完整结构及其与核小体结合的复合物三维结构(图 1i)<sup>[24]</sup>. 酿酒酵母中 INO80 的分子质量为 1.3 Mu, 含有 15 个不同的亚基: Snf2 ATPase、Rvb1、Rvb2、Arp4、Arp5、Arp8、Act1 (actin)、Taf14 (TBP-associated factor14)、Nhp10 (nonhistone protein 10) 以及 Ies1-6. 从电镜密度图上来看, INO80 复合物呈现一个延伸的胚胎状结构. Rvb1/2 复合物位于其头部, 组成一个双环异源十二聚物. Snf2 ATPase 坐落于颈部, 被 Ies2 环绕; Arp5-Ies6 模块位于颈部另一侧; Nhp10 模块处于身体部位、Arp8-Arp4-actin 模块则构成脚的部分. 生化数据显示 Arp5-Ies6 与 Arp8-Arp4-actin 两大 ARP 模块与 Nhp10 模块一起参与与核小体的结合. 虽然作者尝试了 INO80 与核小体结合的负染样品电镜观测, 但由于 INO80 的高度动态, 样品中核小体结合率比例不高, 以及可能存在的核小体被重塑的状态, 最终未能得到很好的三维结构. 但是通过二维分析与生化交联实验可以推出, INO80 采取了与 SWI 及 ISWI 均不相同的结合方式. 它既非包裹, 也非抱住, 也不是单面贴着核小体, 而是像一只手那样的握着核小体: 核小体被置于头、身体和脚组成的凹陷处, 为了更好地抓住它, INO80 的脚部末梢会往回抬起折叠, 从而对核小体形成三面的包围<sup>[24]</sup>.

同一年, 哈佛大学 Andres E. Leschziner 实验室则报道了 INO80 家族另一大类型同样来自酿酒酵母中的 SWR1 复合物本身及其与核小体结合的三维结构(图 1j)<sup>[25]</sup>. SWR1 分子质量为 1 Mu, 含 14 个不同的亚基. 除了 Rvb1/2 外, 可以分为三大模块: 催化亚基 Swr1 模块、Bdf1-Arp4-Act1-Swc4-Yaf9-Swc7 构成的 N 模块、Swc3-Swc2-Arp6-Swc6 构成的 C 模块. Swr1 模块的 N 端部分含有 HAS 结构域, 与 Arps 相互作用; N 模块参与对多重乙酰化核小体的识别与结合; C 模块则结合 H2A.Z/H2B 的二聚体. SWR1 的空间拓扑结构与 INO80 有很大不同. 首先, 相对于长条延伸形的

INO80, SWR1 模块排布的更加紧密. 其次, 不同于 INO80 的双环异源十二聚 Rvb 模块, SWR1 的 Rvb 模块是单环异源六聚复合物. 从整体上看, SWR1 的 C 模块与 N 模块独立存在但是相互挨着, Swr1 模块和 Rvb1/2 模块则如同括号一样围在两边. 不同于 INO80, 作者成功地得到了 SWR1 与核小体结合的复合物结构, 可以明显发现, 加入核小体后, Swr1 模块与 C 模块以远离 Rvb1/2 环的方式延伸. 核小体位于 Swr1 ATPase 与 Rvb1/2 形成

的凹陷处, 一面朝着 SWR1 复合物, 另一面则暴露于外. 核小体与复合物的主要联系由 Swr1 模块提供, 相对于其他的重塑复合物而言, 整体的联系弱了很多<sup>[25]</sup>.

虽然电镜结果展示的都是 INO80 和 SWR1 与单个核小体结合的方式, 但是在生理条件下它们有可能结合更多数量的核小体. 我们期待, 未来冷冻电镜技术的发展, 能带给我们更详细更动态的信息.

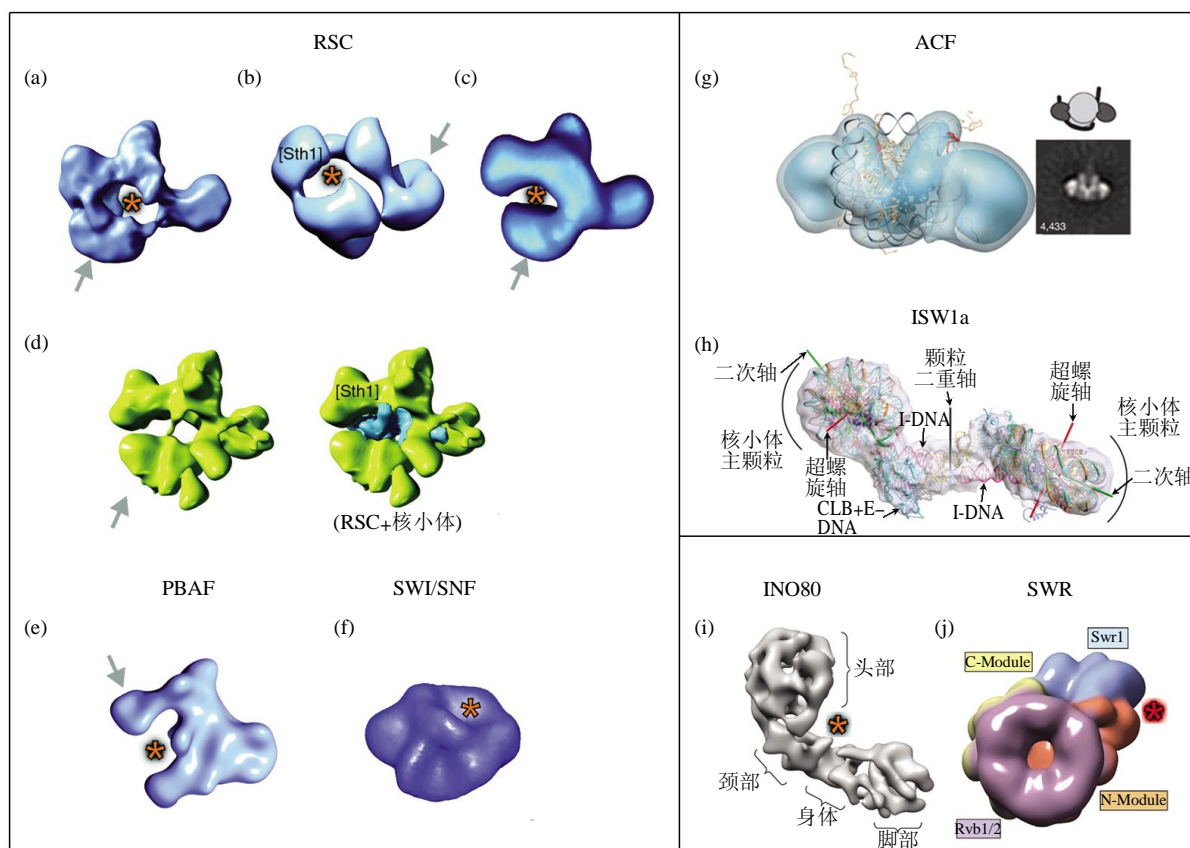


Fig. 1 Electron microscopy reconstructions of ATP-dependent chromatin remodelers

图 1 ATP 依赖的染色质重塑复合物电镜密度图

本图分为 3 个模块, 左边为 SWI 家族, 右上为 ISWI 家族, 右下为 INO80 家族. 左边 SWI 家族模块中(a), (b), (c)为酿酒酵母中 RSC 复合物负染电镜结构<sup>[9,12-13]</sup>, (d) RSC 冷冻电镜结构<sup>[10]</sup>, (e) 人源 PBAF 负染结构<sup>[11]</sup>, (f) 酿酒酵母中 SWI/SNF 结构<sup>[14]</sup>. (a~e)中箭头示复合物中具有柔性的地方. 右上 ISWI 家族模块中(g)为人源 ACF 蛋白中 snf2 亚基与核小体结合的负染色电镜三维重构结构<sup>[20]</sup>, (h) 酿酒酵母 ISW1a 与核小体结合的二聚态冷冻电镜结构<sup>[21]</sup>. 右下 INO80 家族模块中(i)为酿酒酵母 INO80 的负染结构<sup>[24]</sup>, (j) 酿酒酵母中 SWR 负染结构<sup>[25]</sup>. 整个图中星号(\*)标记核小体的可能结合位置.

## 1.2 核小体结合因子

NuA4 是一个含有 13 个不同亚基的组蛋白乙酰转移酶(histone acetyltransferase). 它的催化亚基 Esa1 与其他 2 个亚基 Yng2、Epl1 可以组成一个独立行使催化功能的亚复合物, 命名为 Piccolo-NuA4, 简称 Piccolo. 2011 年 Francisco J Asturias 实验组发

表了 Piccolo 与核小体的 20Å 分辨率水平的冷冻电镜结构, 显示 Piccolo 用 2 个延伸的爪状结构 1:1 与核小体结合, 辅助的 cross-linking 实验将 Esa1 的 Tudor 结构域置于核小体 DNA 处, Epl1 的 EPcA 结构域则在 H2A 的 N 端尾部位置(图 2a)<sup>[26]</sup>.

HP1 (heterochromatin protein 1)是一个非常重要

的表观遗传标记分子, 它有 2 个结构域, chromo domain 和 chromo shadow domain, 由一个没有结构的链状区域(hinge region)连接. HP1 能识别 H3K9 甲基化的核小体, 参与基因沉默、染色质区域化等一系列重要过程. 2013 年 Canzio 等<sup>[27]</sup>报道了 HP1 的同源蛋白 Swi6 二聚体与单个核小体结合并处于非抑制状态的 25Å 电镜结构(图 2b). 从密度图上可以看出, 某一个 Swi6 的 chromo domain 与另一个的 chromo shadow domain 在 H3 尾端的出口处以及 SHL  $\pm 5$  DNA 处与核小体有连接. 另外的 chromo domain 朝向临近核小体的 H3K9me3 位置, 以便于 Swi6 在核小体之间扩散(spreading).

整合子(intasome)是一种四聚体的整合酶(integrase), 介导逆转录病毒将病毒遗传信息整合到宿主 DNA 中的过程, 是与核小体相互作用的重

要因子, 也是当前治疗逆转录病毒如 HIV 的重要靶标. 2015 年, Peter Cherepanov 实验小组成功地使用冷冻电镜单颗粒三维重构技术得到了 PFV 病毒(protofoamy virus)整合子与人类 D02 核小体的复合物三维结构, 分辨率达到了 7.8Å, 是目前已发表的与核小体相关的复合物结构中最高分辨率的结构(图 2c). 借助于整合子本身以及核小体本身的晶体结构, 可以直观地看到整合子本身的三个整合酶亚基, 核小体的两股 DNA 螺旋以及组蛋白上的一个 H2A/H2B 二聚体均参与了整合子与核小体的相互作用. 整合子携带的病毒 DNA 与核小体的 DNA 在该 H2A/H2B 二聚体上侧相遇(远离 Dyad 36 bp 处), 并将核小体的 DNA 带着偏离了其原先缠绕于组蛋白上的位置<sup>[28]</sup>.

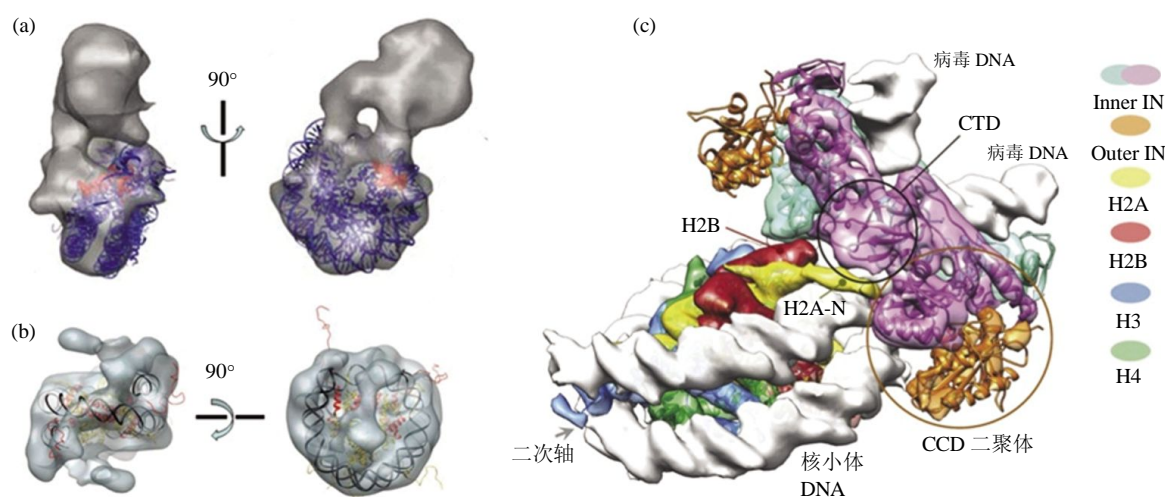


Fig. 2 Cryo-EM reconstruction of nucleosome binding factors with nucleosome

图 2 核小体结合因子与核小体结合的三维电镜密度图

(a) Piccolo-NUA4 与核小体结合冷冻密度图<sup>[26]</sup>. (b) Swi6 与核小体结合冷冻密度图<sup>[27]</sup>. (c) PFV 整合子与核小体结合冷冻密度图<sup>[28]</sup>.

近些年还有一些表观遗传修饰相关的重要催化酶电镜结构工作发表, 如人源 PRC2 复合物及其辅助因子 AEBP2 的负染结构<sup>[29]</sup>, 酵母 Set1/COMPASS 复合物和人源 MLL/COMPASS-like 复合物的电镜结构<sup>[30]</sup>等. 但这些结构的解析分辨率都相对较低, 而且没有获得与核小体复合体的结构, 还有很多工作可以做.

## 2 染色质高级结构的研究

染色质结构在诸多细胞核调控过程中都起到了根本性的作用, 例如 DNA 的转录、复制、重组以

及修复过程. 在过去的 30 多年里, 人们一直试图获得由核小体在连接组蛋白帮助下折叠形成染色质二级结构的高分辨率信息. 通过电子显微镜、X 射线及中子散射法、沉降分析法和 X 射线晶体学等研究手段<sup>[31]</sup>, 研究者提出了一系列关于 30 nm 染色质纤维结构的模型, 主要分为两种: one-start helix 模型和 two-start helix 模型. 其中 one-start helix 模型又称为螺线管(solenoid)模型, 在此模型中, 核小体串珠沿着螺旋轴盘绕上升形成纤维, 一级结构上相邻的核小体在形成的二级结构上也处于相邻的位置, 并且连接 DNA 是弯曲的以适应不同的 DNA



长度<sup>[32]</sup>。由此模型可以推测染色质纤维的直径是由核小体之间固有的相互作用决定,而非核小体重单元的长。而 two-start helix 模型是之字形(zig-zag)模型,这种模型的一个基本特征是在压缩的过程中,连接 DNA 仍处于延伸状态且是直的。它又可以分成交叉连接(crossed-linker)模型和螺旋桨(helical-ribbon)模型,一级序列上相邻的核小体分列在螺旋轴的两侧,连接 DNA 分别垂直和平行于螺旋轴<sup>[33-34]</sup>。相比于螺旋管模型, two-start helix 模型中不同的连接 DNA 长度对形成的染色质纤维直径和长度影响很大。

近年来,随着电镜技术的应用,尤其冷冻电镜技术的应用,关于 30 nm 染色质纤维的结构模型研究,有了一些更深入的进展。根据 30 nm 染色质纤维实验材料的来源,这些研究可以大致分为以下几类:体内提取(isolation)、体内原位观察(observation *in situ*)和体外重组表达(reconstituted *in vitro*)。

## 2.1 体内提取 30 nm 染色质纤维结构研究

人们借助电子显微镜,从各种不同的生物体细胞(海参细胞、鸡红细胞、美洲蝶螈红细胞、鼠肝细胞等)中提取出染色质纤维进行观察研究<sup>[32-35]</sup>。1974 年,Olins 夫妇和 Woodcock 等从真核细胞中提取染色质并在电镜下观察负染及正染样品,从而提出了串珠状结构(beads-on-a-string)这一概念。之后,Finch 和 Klug<sup>[32]</sup>从鼠肝细胞核中提取出天然 11 nm 核小体串珠状结构,并向其中加入连接组蛋白 H1 或  $Mg^{2+}$ ,在电镜下观察到了 30 nm 染色质纤维结构。根据电镜结果,他们还提出了 one-start 的螺旋管模型。

Woodcock 等<sup>[34]</sup>从鸡红细胞中提取出染色质纤维,分别在不同的盐浓度下透析(10 mmol/L NaCl、10 mmol/L  $MgCl_2$ 、20 mmol/L NaCl、20 mmol/L  $MgCl_2$ 、50 mmol/L NaCl、50 mmol/L  $MgCl_2$ 、100 mmol/L NaCl、150 mmol/L KCl),利用透射电镜和扫描电镜进行观察分析。他们发现,随着盐离子浓度的提高,染色质凝聚程度变大,结构也变得更加紧密,其密度由 10 mmol/L NaCl 时的  $6.0 \times 10^4$  u/nm 增加到 150 mmol/L 盐浓度下的  $27 \times 10^4$  u/nm,而压缩比由 2.5 个核小体 / 11 nm 变为 11.6 个核小体 / 11 nm。通过对电镜数据进行整理分析,他们认为染色质纤维符合 two-start 的螺旋桨(helical-ribbon)模型,并推测连接组蛋白 H1 所在的位置都是相同的,并不受连接 DNA 长度的影响,且既不完全位

于纤维外部,也不完全位于隐藏于内部。在之后对长纤维(超 1000 个核小体)的研究中,他们进一步验证了之前的模型,并提出染色质纤维是不规则的 3D zig-zag 模型<sup>[36]</sup>。Langmore 等<sup>[37]</sup>用海参精子细胞和蝶螈红细胞作为实验材料提取染色质,在冷冻状态下观察,得出与 Woodcock 不同的结论。他们提出染色质呈现 two-start 的交叉连接(crossed-linker)模型,双螺旋纤维组成的核小体呈现“face to face”的排列方式,两列核小体相互平行且平行于染色质纤维螺旋轴。Langmore 等还发现随着离子浓度的升高,zig-zag 结构会变得更加紧密,这与 Woodcock 结论相同。同时,染色质纤维的折叠与去折叠过程非常快,时间不到 10 ms<sup>[33]</sup>。

最近,Frangakis 团队利用冷冻电子断层成像技术,对从鸡红细胞中提取的染色质纤维进行冷冻观察,发现在紧实的染色质纤维中,核小体主要是以“face to face”的方式堆叠排布。他们推测核小体的堆积叠加是体内染色质高级结构形成的重要机制<sup>[38]</sup>。

## 2.2 体内原位观察 30 nm 染色质纤维结构研究

利用体内原位观察的染色质研究成果也有很多,其优势在于可以真实反映分裂间期染色质的形态结构。早期,英国的 Davies 等<sup>[39-40]</sup>利用电子显微镜观察不同时期的鸡红细胞切片,实现了对细胞内完整染色质的初步观察研究。Langmore 等<sup>[37]</sup>利用电镜观察鸡红细胞核普通包埋切片,发现存在“side-by-side”排列的直径在 20~30 nm 之间的纤维,确认了鸡红细胞中 30 nm 染色质纤维的存在。Mcdowall 及其合作者则对中国仓鼠卵巢细胞和 HeLa 细胞冷冻切片进行观察,发现染色质呈现出有特征、均一、有颗粒状的纹理,但是并没有观察到更高级的规律性结构<sup>[41]</sup>。Woodcock 团队对海星精子细胞、鸡红细胞核、海蝙蝠精子细胞、海参精子细胞等多个生物物种细胞内的染色质进行电镜观察,发现含有不同长度连接 DNA 的染色质,均能看到明显的规律性存在的染色质高级结构。他们还利用电子断层成像技术对海星精子头部细胞进行了详细的分析,发现其体内的染色质呈现不对称的 zig-zag 排列,三维构象有无序的弯曲和扭转<sup>[41-43]</sup>。

Frangakis 团队针对 HeLa 细胞和鸡红细胞中的染色质进行了大量的研究,他们利用高压冷冻的办法固定 HeLa S3 细胞,并制成冷冻切片以尽量接近天然状态,得到高分辨率的成像。在对图像进行处理分析后他们发现,在有丝分裂染色体中,染色

质纤维处在高度无序、相间错杂的状态, 并且随着溶液中镁离子的减少, 染色质会发生膨胀, 从而抑制 30 nm 染色质纤维紧密结构的形成. 由此, 他们提出有丝分裂染色体中没有发现 30 nm 染色质纤维存在<sup>[44]</sup>. 之后, Mikhail 等也提出在有丝分裂期染色体中可能并不存在 30 nm 染色质结构, 并提出熔融模型(melting model)的假说, 即体内的染色体是处于一种高度无序, 相间交错的状态, 类似于多聚体熔融的动态移动. 在体外构建系统中, 核小体浓度低, 纤维之间的相互作用可以忽略不计, 其内部核小体的相互作用会导致 30 nm 染色质纤维结构的形成. 而在体内核小体浓度很高, 纤维之间核小体的相互作用逐渐占据主导地位, 于是 30 nm 结构“溶解”成了尺度较小的无规则动态结构. 他们还假设在细胞间期染色体也处于类似的状态, 从而有助于转录过程和 DNA 复制过程的进行<sup>[45]</sup>.

Frangakis 团队<sup>[46]</sup>对鸡血红细胞核冷冻切片又进行了分析, 利用冷冻电子断层成像三维重构后, 三维分类的结构显示天然染色质是 two-start 左手螺旋的构象, 沿着螺旋轴方向每 11 nm 的距离约 6.5 个核小体, 核小体呈“face to face”并列排布. 从垂直于轴的方向观察, 核小体呈现棋盘式分布, 而不是相间错落的形式, 并且核小体之间“中心对中心”的距离比预期的大, 可能是为了给转录和调控因子接近核小体预留空间.

### 2.3 体外重组 30 nm 染色质纤维结构研究

虽然利用细胞内提取的天然染色质或者体内原位观察进行结构研究更能了解真实生理状态下的染色质结构形态, 但是由于 DNA 序列、组蛋白成分、组蛋白翻译后修饰的不同, 以及核小体之间相对取向的不规则性, 使得天然染色质具有高度的异质性和可变性, 从而导致利用电子显微镜无法获得高分辨率的二级结构模型. 为了获得尽量均一的研究材料, 人们建立了一套体外重构组装的体系, 即: 将体外表达的组蛋白八聚体与一种人工设计的 601 DNA<sup>[47]</sup>(含有 wisdom 601 nucleosome positioning 序列, 对组蛋白八聚体有很强的亲和力)按一定比例混合, 经过透析形成串珠状结构, 之后加入二价阳离子(一般是  $Mg^{2+}$ )或连接组蛋白 H1/H5, 形成染色质纤维二级结构. 在组装过程中, 可以通过改变不同的条件, 如组蛋白修饰、组蛋白变体、连接 DNA 的长度以及透析环境等获得不同状态的纤维结构, 有助于探究表观遗传调控过程中染色质的动态变化机制.

Richmond 研究组用 48×177 bp DNA 长度的染色质作为实验材料制成负染样品, 在电镜下观察发现大多数纤维都是宽度在 25~30 nm 之间, 等长的双排平行结构, 看上去与 two-start 模型相符. 他们还观察了 10×208 bp DNA 长度染色质的负染样品, 在组装的过程中按照核小体与连接组蛋白 1:1 的比例加入 H1, 结果在电镜下也是看到平行的双排核小体, 由此得出连接组蛋白结合核小体不会影响染色质纤维二级结构是 one-start 还是 two-start<sup>[48]</sup>. 英国剑桥大学的 Rhodes 研究组为了研究连接 DNA 长度对纤维的影响, 体外构建了一系列不同 DNA 长度的染色质纤维, 其中包括 72×177 bp、52×187 bp、61×197 bp、47×207 bp、55×217 bp、66×227 bp 以及 56×237 bp, 并在染色质折叠过程中加入连接组蛋白 H5 和 1.6 mmol/L 的  $MgCl_2$ . 利用电子显微镜观察负染以及冷冻样品并测量纤维直径后, 他们得出结论: 染色质纤维的直径并没有随着连接 DNA 长度的增加而呈线性增长, 这与 two-start 模型是相悖的. 当连接 DNA 长度在 10~40 bp 时, 染色质纤维直径为 33 nm, 核小体密度是每 11 nm 的距离内含有 11 个核小体. 而当连接 DNA 在 50~70 bp 时, 染色质纤维的直径变为 44 nm, 每 11 nm 距离内含有 15 个核小体. Rhodes 等则根据实验数据提出了一种左手 one-start 交叉螺旋管模型, 并推测连接组蛋白会通过改变连接 DNA 的轨迹从而决定连续核小体之间的位置以及相互作用模式, 最终决定核小体串珠结构折叠成染色质时形成的形状. 同时, 他们也提出, 关于染色质结构更详细的信息需要通过获得高分辨率的结果才能得以阐释<sup>[49]</sup>.

最近, 中国科学院生物物理研究所多个课题组通力合作, 利用冷冻电镜单颗粒三维重构技术, 解析得到了体外组装的由 12 个和 24 个核小体组成的 30 nm 染色质纤维高分辨率(11 Å)结构<sup>[49]</sup>. 该结构清晰展示(图 3a): 体外组装、由连接组蛋白 H1 介导形成的 30 nm 染色质纤维会形成一个左手双螺旋的 two-start 结构. 在这个染色质左手双螺旋结构模型中, 每 4 个核小体形成一结构单元; 结构单元和单元之间的相互扭曲形成染色质双螺旋结构; 同时, 四聚核小体单元之间有空隙, 这个空隙可能是组蛋白修饰、染色质重塑等重要表观遗传现象发生的调控区域. 同时, 该结构还揭示(图 3b)连接组蛋白 H1 在单个核小体内部以及核小体单元之间呈不对称分布, 这些不对称相互作用促成了 30 nm 高级结

构的形成, 首次明确了 H1 在 30 nm 染色质纤维形成过程中的重要作用. 如上所述, 细胞核中染色质的整体组织形式除了受到细胞周期调控以外, 还会受到各种表观遗传因素(如组蛋白变体、DNA 和组蛋白化学修饰等)的影响. 尽管如此, 这个研究中

所揭示的在体外、低盐条件、依赖于连接组蛋白 H1 组装形成的 30 nm 染色质左手双螺旋高级结构, 主要是由于 DNA 和组蛋白的内在物理和化学性质决定的, 其基本原则对体内染色质也可能适用.

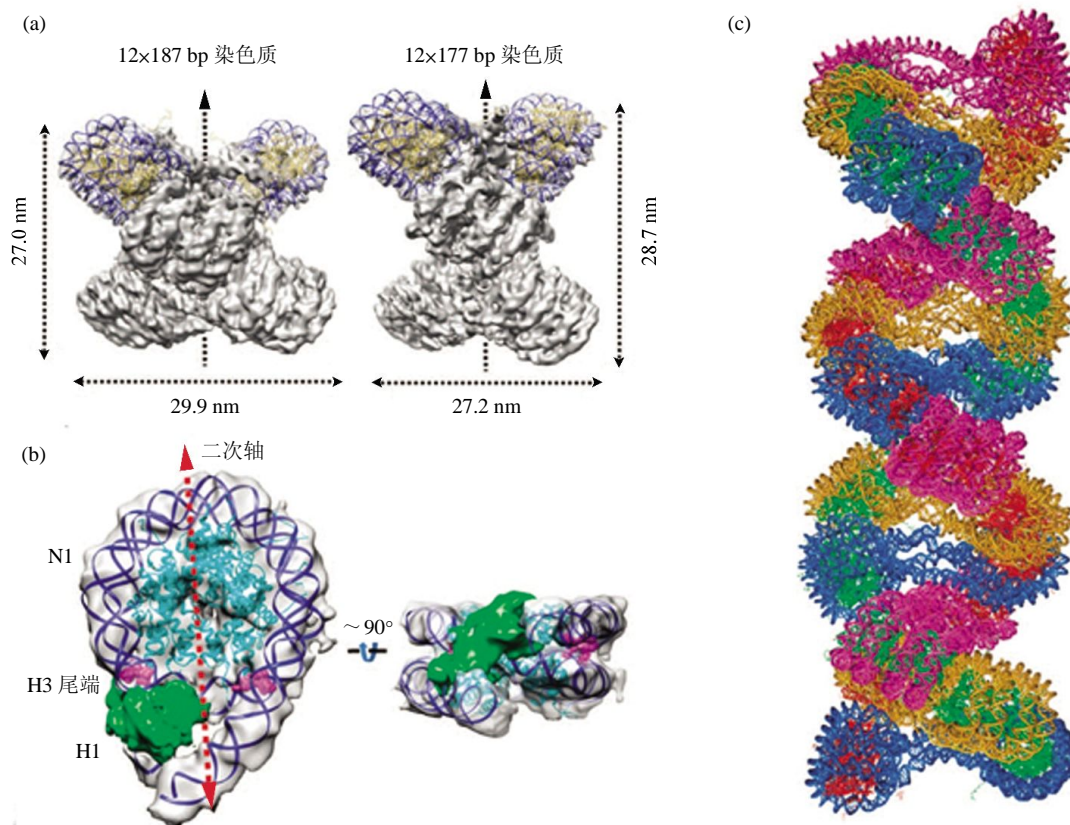


Fig. 3 Cryo-EM reconstruction and pseudo-atomic model of the 30 nm chromatin fiber

图3 体外组装的 30 nm 染色质纤维电镜结构及原子模型

(a) 12x187 bp 与 12x177 bp 染色质纤维三维冷冻电镜密度图. (b) H1 与核小体的非对称结合. (c) 30 nm 染色质纤维的假原子模型<sup>[49]</sup>.

### 3 结 语

随着冷冻电镜技术在硬件与软件上的双向突破, 越来越多的原子水平分辨率的电镜结构得到解析. 而以核小体为核心的表观遗传世界里的电镜故事, 现在才刚刚开启序幕. 在过去 30 多年的研究中, 虽然人们对于染色质与表观遗传调控机制之间的关系有了很多的认识, 但是对于 30 nm 染色质纤维本身的高级结构及天然染色质的结构模型仍然争议颇多. 冷冻电子显微技术作为研究核小体-结构因子复合物、30 nm 染色质纤维等超大分子复合体的重要手段之一, 潜能巨大. 有理由相信, 未来几年会有大量的与核小体、染色质及表观遗传调控相

关复合物的高分辨高难度结构问世, 帮助我们更本质更精细的角度去理解表观遗传的原理与机制, 更加震撼与深刻地去认识生命世界的精彩与神奇.

### 参 考 文 献

- [1] Bai X C, McMullan G, Scheres S H. How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci*, 2015, **40**(1): 49–57
- [2] McGinty R K, Tan S. Nucleosome structure and function. *Chem Rev*, 2015, **115**(6): 2255–2273
- [3] Cairns B R. Chromatin remodeling: insights and intrigue from single-molecule studies. *Nat Struct Mol Biol*, 2007, **14**(11): 989–996
- [4] Singleton M R, Dillingham M S, Wigley D B. Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev*



- Biochem, 2007, **76**: 23–50
- [5] Cote J, Quinn J, Workman J L, *et al.* Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex. *Science*, 1994, **265**(5168): 53–60
- [6] Mujtaba S, Zeng L, Zhou M M. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene*, 2007, **26**(37): 5521–5527
- [7] Kasten M M, Clapier C R, Cairns B R. SnapShot: Chromatin remodeling: SWI/SNF. *Cell*, 2011, **144**(2): 310–e311
- [8] Leschziner A E. Electron microscopy studies of nucleosome remodelers. *Curr Opin Struct Biol*, 2011, **21**(6): 709–718
- [9] Asturias F J, Chung W H, Kornberg R D, *et al.* Structural analysis of the RSC chromatin-remodeling complex. *Proc Natl Acad Sci USA*, 2002, **99**(21): 13477–13480
- [10] Chaban Y, Ezeokonkwo C, Chung W H, *et al.* Structure of a RSC-nucleosome complex and insights into chromatin remodeling. *Nat Struct Mol Biol*, 2008, **15**(12): 1272–1277
- [11] Leschziner A E, Lemon B, Tjian R, *et al.* Structural studies of the human PBAF chromatin-remodeling complex. *Structure*, 2005, **13**(2): 267–275
- [12] Leschziner A E, Saha A, Wittmeyer J, *et al.* Conformational flexibility in the chromatin remodeler RSC observed by electron microscopy and the orthogonal tilt reconstruction method. *Proc Natl Acad Sci USA*, 2007, **104**(12): 4913–4918
- [13] Skiniotis G, Moazed D, Walz T. Acetylated histone tail peptides induce structural rearrangements in the RSC chromatin remodeling complex. *J Biol Chem*, 2007, **282**(29): 20804–20808
- [14] Smith C L, Horowitz-Scherer R, Flanagan J F, *et al.* Structural analysis of the yeast SWI/SNF chromatin remodeling complex. *Nat Struct Biol*, 2003, **10**(2): 141–145
- [15] Dechassa M L, Zhang B, Horowitz-Scherer R, *et al.* Architecture of the SWI/SNF-nucleosome complex. *Mol Cell Biol*, 2008, **28**(19): 6010–6021
- [16] Yadon A N, Tsukiyama T. SnapShot: Chromatin remodeling: ISWI. *Cell*, 2011, **144**(3): 453–453 e451
- [17] Grune T, Brzeski J, Eberharder A, *et al.* Crystal structure and functional analysis of a nucleosome recognition module of the remodeling factor ISWI. *Mol Cell*, 2003, **12**(2): 449–460
- [18] Dang W, Bartholomew B. Domain architecture of the catalytic subunit in the ISW2-nucleosome complex. *Mol Cell Biol*, 2007, **27**(23): 8306–8317
- [19] Hu M, Zhang Y B, Qian L, *et al.* Three-dimensional structure of human chromatin accessibility complex hCHRAC by electron microscopy. *J Struct Biol*, 2008, **164**(3): 263–269
- [20] Racki L R, Yang J G, Naber N, *et al.* The chromatin remodeller ACF acts as a dimeric motor to space nucleosomes. *Nature*, 2009, **462**(7276): 1016–1021
- [21] Yamada K, Frouws T D, Angst B, *et al.* Structure and mechanism of the chromatin remodelling factor ISWIa. *Nature*, 2011, **472**(7344): 448–453
- [22] Chen L, Cai Y, Jin J, *et al.* Subunit organization of the human INO80 chromatin remodeling complex: an evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodeling. *J Biol Chem*, 2011, **286**(13): 11283–11289
- [23] Gerhold C B, Gasser S M. INO80 and SWR complexes: relating structure to function in chromatin remodeling. *Trends Cell Biol*, 2014, **24**(11): 619–631
- [24] Tosi A, Haas C, Herzog F, *et al.* Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell*, 2013, **154**(6): 1207–1219
- [25] Nguyen V Q, Ranjan A, Stengel F, *et al.* Molecular architecture of the ATP-dependent chromatin-remodeling complex SWR1. *Cell*, 2013, **154**(6): 1220–1231
- [26] Chittuluru J R, Chaban Y, Monnet-Saksouk J, *et al.* Structure and nucleosome interaction of the yeast NuA4 and Piccolo-NuA4 histone acetyltransferase complexes. *Nat Struct Mol Biol*, 2011, **18**(11): 1196–1203
- [27] Canzio D, Liao M, Naber N, *et al.* A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. *Nature*, 2013, **496**(7445): 377–381
- [28] Maskell D P, Renault L, Serrao E, *et al.* Structural basis for retroviral integration into nucleosomes. *Nature*, 2015, **523**(7560): 366–369
- [29] Ciferri C, Lander G C, Maiolica A, *et al.* Molecular architecture of human polycomb repressive complex 2. *Elife*, 2012, **1**: e00005
- [30] Takahashi Y H, Westfield G H, Oleskie A N, *et al.* Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. *Proc Natl Acad Sci USA*, 2011, **108**(51): 20526–20531
- [31] Robinson P J, Fairall L, Huynh V A, *et al.* EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proc Natl Acad Sci USA*, 2006, **103**(17): 6506–6511
- [32] Finch J T, Klug A. Solenoidal model for superstructure in chromatin. *Proc Natl Acad Sci USA*, 1976, **73**(6): 1897–1901
- [33] Athey B D, Smith M F, Rankert D A, *et al.* The diameters of frozen-hydrated chromatin fibers increase with DNA linker length: evidence in support of variable diameter models for chromatin. *J Cell Biol*, 1990, **111**(3): 795–806
- [34] Woodcock C L, Frado L L, Rattner J B. The higher-order structure of chromatin: evidence for a helical ribbon arrangement. *J Cell Biol*, 1984, **99**(1 Pt 1): 42–52
- [35] Daban J R. Electron microscopy and atomic force microscopy studies of chromatin and metaphase chromosome structure. *Micron*, 2011, **42**(8): 733–750
- [36] Horowitz R A, Koster A J, Walz J, *et al.* Automated electron microscope tomography of frozen-hydrated chromatin: the irregular three-dimensional zigzag architecture persists in compact, isolated fibers. *J Struct Biol*, 1997, **120**(3): 353–362
- [37] Langmore J P, Paulson J R. Low angle x-ray diffraction studies of chromatin structure *in vivo* and in isolated nuclei and metaphase chromosomes. *J Cell Biol*, 1983, **96**(4): 1120–1131
- [38] Scheffer M P, Eltsov M, Bednar J, *et al.* Nucleosomes stacked with aligned dyad axes are found in native compact chromatin *in vitro*.

- J Struct Biol, 2012, **178**(2): 207–214
- [39] Everid A C, Small J V, Davies H G. Electron-microscope observations on the structure of condensed chromatin: evidence for orderly arrays of unit threads on the surface of chicken erythrocyte nuclei. J Cell Sci, 1970, **7**(1): 35–48
- [40] Davies H G, Murray A B, Walmsley M E. Electron-microscope observations on the organization of the nucleus in chicken erythrocytes and a superunit thread hypothesis for chromosome structure. J Cell Sci, 1974, **16**(2): 261–299
- [41] McDowell A W, Smith J M, Dubochet J. Cryo-electron microscopy of vitrified chromosomes *in situ*. EMBO J, 1986, **5**(6): 1395–1402
- [42] Woodcock C L. Chromatin fibers observed *in situ* in frozen hydrated sections. Native fiber diameter is not correlated with nucleosome repeat length. J Cell Biol, 1994, **125**(1): 11–19
- [43] Horowitz R A, Agard D A, Sedat J W, *et al.* The three-dimensional architecture of chromatin *in situ*: electron tomography reveals fibers composed of a continuously variable zig-zag nucleosomal ribbon. J Cell Biol, 1994, **125**(1): 1–10
- [44] Eltsov M, MacLellan K M, Maeshima K, *et al.* Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes *in situ*. Proc Natl Acad Sci USA, 2008, **105**(50): 19732–19737
- [45] Maeshima K, Hihara S, Eltsov M. Chromatin structure: does the 30-nm fibre exist *in vivo*? Curr Opin Cell Biol, 2010, **22** (3): 291–297
- [46] Scheffer M P, Eltsov M, Frangakis A S. Evidence for short-range helical order in the 30-nm chromatin fibers of erythrocyte nuclei. Proc Natl Acad Sci USA, 2011, **108**(41): 16992–16997
- [47] Lowary P T, Widom J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J Mol Biol, 1998, **276**(1): 19–42
- [48] Dorigo B, Schalch T, Kulangara A, *et al.* Nucleosome arrays reveal the two-start organization of the chromatin fiber. Science, 2004, **306**(5701): 1571–1573
- [49] Song F, Chen P, Sun D, *et al.* Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. Science, 2014, **344**(6182): 376–380

## Application of Cryo-Electron Microscopy on Epigenetic Studies\*

LI Cheng-Min<sup>1,2)\*\*</sup>, ZHOU Min<sup>1,2)\*\*</sup>, LI Xiao-Min<sup>1,2)</sup>, ZHU Ping<sup>1)\*\*\*</sup>

<sup>(1)</sup> National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China;

<sup>(2)</sup> University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract** The architecture and structural details of chromatin assembler, chromatin modifier, and chromatin remodeler, and their complexes with nucleosome and chromatin, are critical for revealing the molecular mechanisms for the establishment, maintenance and regulation of epigenetic information. The fast-developing cryo-electron microscopy (cryo-EM) provides an indispensable tool for the structure determination of the huge, multi-subunits, and flexible epigenetic regulation related macromolecular assemblies. In this review, we summarized the recent progresses on EM structural studies of macromolecules and complexes relating to epigenetic regulation.

**Key words** epigenetic regulation, cryo-EM, nucleosome, chromatin

**DOI:** 10.16476/j.pibb.2015.0295

\* This work was supported by grants from National Basic Research Program of China (2015CB856200), Strategic Priority Research Program (XDB08010100) of the Chinese Academy of Sciences, Key Research Program (KJZD-EW-L05) from The Chinese Academy of Sciences, and The National Natural Science Foundation of China (31230018, 31425007, 31521002).

\*\*These authors contributed equally to this work.

\*\*\*Corresponding author.

Tel: 86-10-64888799, E-mail: zhup@ibp.ac.cn

Received: September 25, 2015 Accepted: September 28, 2015